



Méthodes de Fusion de Cartes de Caractéristiques pour la Détection Multispectrale par Réseaux de Neurones Profonds

Heng Zhang, Elisa Fromont, Sébastien Lefèvre, Bruno Avignon

► To cite this version:

Heng Zhang, Elisa Fromont, Sébastien Lefèvre, Bruno Avignon. Méthodes de Fusion de Cartes de Caractéristiques pour la Détection Multispectrale par Réseaux de Neurones Profonds. RFIAP 2020 - Reconnaissance des Formes, Image, Apprentissage et Perception, Jun 2020, Vannes, France. pp.1-2. hal-02872123

HAL Id: hal-02872123

<https://hal.science/hal-02872123>

Submitted on 17 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Méthodes de Fusion de Cartes de Caractéristiques pour la Détection Multispectrale par Réseaux de Neurones Profonds

Heng ZHANG^{1,3}

Elisa FROMONT^{1,4}

Sébastien LEFEVRE²

Bruno AVIGNON³

¹ Univ Rennes, IRISA ² Univ Bretagne Sud, IRISA ³ ATERMES ⁴ IUF, Inria

heng.zhang@irisa.fr

Résumé

L'utilisation d'images multispectrales (par exemple des paires d'images en bandes spectrales visible et infrarouge) peut s'avérer particulièrement utile lorsque l'on cherche à détecter des objets dans des environnements variés (par exemple des scènes extérieures capturées de jour ou de nuit). Pour utiliser ces différentes bandes spectrales, le principal problème technique est la fusion des informations complémentaires issues des différentes bandes. Si cette fusion peut, en théorie, être mise en oeuvre à différents niveaux (précoce, intermédiaire, tardive), les méthodes récentes d'apprentissage profond ont montré que la fusion à mi-parcours (intermédiaire) dans le réseau était celle donnant le meilleur gain de performances. Cet article propose deux nouvelles approches intitulées PS-Fuse et Cyclic Fuse-and-Refine pour fusionner au mieux les caractéristiques multispectrales au sein d'un réseau de neurones profond. Nos expériences montrent que ces deux contributions conduisent à des améliorations significatives de précision par rapport aux méthodes existantes en détection d'objets sur des jeux de données multispectrales.

Mots clés

Détection d'objets sur des images multi-spectrales, fusion multi-spectrale, apprentissage profond.

Références

- [1] Soonmin Hwang, Jaesik Park, Namil Kim, Yukyung Choi, and In So Kweon. Multispectral pedestrian detection : Benchmark dataset and baselines. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [2] Chengyang Li, Dan Song, Ruofeng Tong, and Min Tang. Illumination-aware faster R-CNN for robust multispectral pedestrian detection. *Pattern Recognition*, 85 :161–171, 2019.
- [3] Dayan Guan, Yanpeng Cao, Jiangxin Yang, Yanlong Cao, and Michael Ying Yang. Fusion of multispectral data through illumination-aware deep neural networks for pedestrian detection. *Information Fusion*, 50 :148–157, 2019.
- [4] Chengyang Li, Dan Song, Ruofeng Tong, and Min Tang. Multispectral pedestrian detection via simultaneous detection and segmentation. In *British Machine Vision Conference (BMVC)*, page 225, 2018.
- [5] Lu Zhang, Zhiyong Liu, Shifeng Zhang, Xu Yang, Hong Qiao, Kaizhu Huang, and Amir Hussain. Cross-modality interactive attention network for multispectral pedestrian detection. *Information Fusion*, 50 :20–29, 2019.
- [6] Lu Zhang, Xiangyu Zhu, Xiangyu Chen, Xu Yang, Zhen Lei, and Zhiyong Liu. Weakly aligned cross-modal learning for multispectral pedestrian detection. In *IEEE International Conference on Computer Vision (ICCV)*, 2019.
- [7] Jingjing Liu, Shaoting Zhang, Shu Wang, and Dimitris N. Metaxas. Multispectral deep neural networks for pedestrian detection. In *Proceedings of the British Machine Vision Conference 2016, BMVC 2016, York, UK, September 19-22, 2016*, 2016.
- [8] Free flir thermal dataset for algorithm training. <https://www.flir.com/oem/adas/adas-dataset-form/>.
- [9] Jörg Wagner, Volker Fischer, Michael Herman, and Sven Behnke. Multispectral pedestrian detection using deep fusion convolutional neural networks. In *24th European Symposium on Artificial Neural Networks (ESANN)*, 2016.
- [10] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN : towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 91–99, 2015.
- [11] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. arXiv :1312.4400, 2013.
- [12] Lu Zhang, Zhiyong Liu, Shifeng Zhang, Xu Yang, Hong Qiao, Kaizhu Huang, and Amir Hussain. Cross-modality interactive attention network for multispectral pedestrian detection. *Information Fusion*, 50 :20–29, 2019.

- [13] Yang Zheng, Izzat H. Izzat, and Shahrzad Ziaee. GFD-SSD : gated fusion double SSD for multispectral pedestrian detection. *CoRR*, abs/1903.06999, 2019.
- [14] Garrick Brazil, Xi Yin, and Xiaoming Liu. Illuminating pedestrians via simultaneous detection & segmentation. In *Proceedings of the IEEE International Conference on Computer Vision (CVPR)*, 2017.
- [15] Xiaohan Ding, Yuchen Guo, Guiguang Ding, and Jungong Han. Acnet : Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks. In *The IEEE International Conference on Computer Vision (ICCV)*, 2019.
- [16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet : A Large-Scale Hierarchical Image Database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [17] Zuoxin Li and Fuqiang Zhou. FSSD : feature fusion single shot multibox detector. *CoRR*, abs/1712.00960, 2017.
- [18] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. Ssd : Single shot multibox detector. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016.
- [19] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.
- [20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers : Surpassing human-level performance on imagenet classification. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1026–1034, 2015.
- [21] Liliang Zhang, Liang Lin, Xiaodan Liang, and Kaiming He. Is faster r-cnn doing well for pedestrian detection ? arXiv :1607.07032, 2016.
- [22] Daniel König, Michael Adam, Christian Jarvers, Georg Layher, Heiko Neumann, and Michael Teutsch. Fully convolutional region proposal networks for multispectral person detection. In *IEEE*.
- [23] Lee R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3) :297–302, 1945.